Effects of spatial variability in light use efficiency on satellite-based NPP monitoring

David P. Turner
Stith T. Gower
Warren B. Cohen
Matthew Gregory
Tom K. Maiersperger

*Department of Forest Science, Oregon State University, Corvallis, OR 97331-7501, USA

Department of Forest Ecology and Management, University of Wisconsin, Madison, WI 53706, USA

USDA PNW Research Station, Corvallis, OR 97331-7501, USA

Received 24 April 2001; received in revised form 23 August 2001; accepted 31 August 2001

Abstract

Light use efficiency (LUE) algorithms are a potentially effective approach to monitoring global net primary production (NPP) using satellite-based sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS). However, these algorithms are applied at relatively coarse spatial resolutions (≥1 km), which may subsume significant heterogeneity in vegetation LUE (εs, g MJ−1) and, hence, introduce error. To examine the effects of spatial heterogeneity on a LUE algorithm, imagery from the Advanced Very High Resolution Radiometer (AVHRR) at ∼1-km resolution was used to implement a LUE approach for NPP estimation over a 25-km² area of corn (Zea mays L.) and soybean (Glycine max Merr) in central Illinois, USA. Results from several εs formulations were compared with a NPP reference surface based on measured NPPs and a high spatial resolution land cover surface derived from Landsat ETM+. Determination of εs based on measurements of biomass production and monitoring of absorbed photosynthetically active radiation (APAR) revealed that εs of soybean was 68% of that for corn. When a LUE algorithm for estimating NPP was implemented in the study area using the assumption of homogeneous cropland and the εs for corn, the estimate for total biomass production was 126% of that from the NPP reference surface. Because of counteracting errors, total biomass production using the soybean εs was closer (86%) to that from the NPP reference surface. Retention of high spatial resolution land cover to assign εs resulted in a total NPP very similar to the reference NPP because differences in leaf phenology between the crop types were small except early in the growing season. These results suggest several alternative approaches to accounting for land cover heterogeneity in εs when implementing LUE algorithms at coarse resolution. © 2002 Elsevier Science Inc. All rights reserved.

1. Introduction

Monitoring global terrestrial net primary production (NPP) is relevant to understanding the global carbon cycle and evaluating effects of interannual climate variation on food and fiber production (Running et al., 1999). One approach to monitoring NPP employs satellite-borne sensors that achieve daily coverage over the Earth’s surface at spatial resolutions of 250–1000 m, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) (Justice et al., 1998). Surface reflectances derived from these sensors are used to infer the fraction of incoming photosynthetically active radiation absorbed by the vegetation (fAPAR) (Asrar, Fuchs, Kaemasu, & Hatfield, 1984; Goetz & Prince, 1996). The combination of fAPAR and an estimate of incident PAR from systems such as the Geostationary Operational Environmental Satellite (Gu & Smith, 1997) can provide an estimate of PAR absorbed by the canopy (APAR). Knowledge of the efficiency with which vegetation converts PAR into biomass (εs) then permits an estimate of NPP. Light use efficiency (LUE) models have been applied regionally (Goetz et al., 1999) and globally (Ruimy, Saugier, & Dedieu, 1994), generally with limited validation. One potential source of error in these implementations is subgrid-scale heterogeneity in εs (Behrenfeld et al., 2001).

Initial studies with crop species suggested that εs was stable across species (Montieth, 1972) and evolutionary...
theory supports the convergence of \( \varepsilon_n \) across vegetation types (Field, 1991). However, as \( \varepsilon_n \) was measured for a wider variety of species and vegetation types, the earlier generalization gave way to recognition of significant variation (Goetz & Prince, 1999; Gower, Kucharik, & Normun, 1999; Ruimy et al., 1994). Climatic constraints on NPP have also been studied and have indicated greater consistency in \( \varepsilon_n \) if APAR is excluded from the annual sum when conditions are unfavorable for production (e.g., Runyon, Waring, Goward, & Welles, 1994). Gross primary production (GPP), the net effect of carboxylation and photorespiration, has been proposed as an alternative basis for LUE algorithms because light absorption is more directly linked to GPP than NPP, i.e., plant species differ in their allocation to autotrophic respiration (Goetz & Prince, 1999). Differences among species in foliar nitrogen concentration, which is linked to photosynthetic potential (Field & Mooney, 1986; Sinclair & Horie, 1989), and in carboxylation biochemistry—notably, the \( C_3 \) and \( C_4 \) pathways—also suggest associated differences in GPP production efficiency (\( \varepsilon_g \)).

To whatever degree that \( \varepsilon_n \) or \( \varepsilon_g \) do vary among vegetation cover types, it becomes important to understand effects of spatial heterogeneity in these production efficiencies on the implementation of LUE algorithms. Stands of aspen and black spruce are commonly found in close association in boreal forests, yet, they have quite different \( \varepsilon_n \)'s (Goetz & Prince, 1996; Gower et al., 1999). \( C_3 \) and \( C_4 \) crop species are often used in rotation (e.g., wheat and lentils) and, hence, are commonly found in adjoining fields, but the crop types differ in \( \varepsilon_n \) (Gower et al., 1999). In these cases, assignment of \( \varepsilon_n \) to a coarse resolution grid cell, which includes both cover types, is problematic.

We used measurements of NPP, incident PAR, and leaf area index (LAI) to evaluate \( \varepsilon_n \) in adjacent corn and soybean fields in central Illinois. We also employed fine-resolution satellite imagery to assess spatial heterogeneity in crop cover type over a 25-km\(^2\) area and develop a NPP reference data layer. Potential effects of implementing a coarse resolution LUE algorithm on NPP estimates in the area were then assessed using imagery from the Advanced Very High Resolution Radiometer (AVHRR) system.

2. Methods

2.1. Land cover

The study was performed on a 5 \times 5-km area of cropland in central Illinois during 1999. The site is part of the network of eddy covariance flux towers associated with AmeriFlux (2001) and the network of Core Validation Sites associated with the MODIS Land Team (2001).

A Landsat ETM+ image dating from July 29, 1999 was acquired from the MODIS Land Team (2001) website to produce the land cover map. Source image characteristics include level 1G processing, a cell size of 30 m, and UTM (WGS84) projection. The image was subsequently resampled to 25-m resolution using a cubic convolution algorithm to match the grain size of supporting field measurements. The cosine of the solar zenith angle (COST) radiometric correction model (Chavez, 1996) was applied to convert the source image digital counts to reflectance. Since pixels falling in deep cloud shadow were found to be the darkest scene element, they were used as dark objects in the radiometric calibration algorithm. They were assumed to have 2% reflectance based on published estimates of shadowed vegetation reflectance (Adams et al., 1995; Hall, Shumabukuro, & Huemmrich, 1995). The positional accuracy of the image was digitally assessed by direct comparison with USGS digital orthoimage quadrangles (DOQs) in the study area.

Land cover mapping was performed using an unsupervised clustering of the six ETM+ reflectance bands. A K-means approach was used, with 30 clusters being initialized along the principal axis of variation within the six-dimensional reflectance space. The algorithm produced a set of 30 stable clusters, with a 98% convergence rate over the course of 20 iterations. Clusters were assigned to five classes: water, urban and built, barren and sparsely vegetated, corn, and soybean. This assignment was performed with reference to the DOQs, air photos, interpreter knowledge, and spectral characteristics examined in bivariate frequency distributions. Some confusion was apparent in the cluster map, most notably between the urban and built and barren and sparsely vegetated categories. Additionally, minor confusion within crop types occurred, for example, where localized drainage features reduced the separability of soybean from corn in more moist (lower reflectance) areas. Given that the spatial boundaries between confused land cover classes was readily apparent, hand digitizing and recoding was performed to eliminate the majority of spectral confusion among classes. For the purposes of this analysis, the urban and built, barren and sparsely settled, and water classes were aggregated to an “Other” class (Fig. 1).

Validation of the classification was based on the 80 points sampled for LAI and NPP (see below) plus 20 additional points distributed randomly over the 5 \times 5-km area. Locations for all land cover validation points were registered to within several meters using an Ashtech GG-24 Surveyor Global Positioning System unit (Ashtech, 2001).

2.2. Net primary production

As part of a related study that compared plot-level NPP measurements to measurements of net ecosystem exchange from an eddy covariance flux tower (Reich, Turner, & Bolstad, 1999), 80 plots were established within a 1-km radius of a centrally located eddy covariance flux tower. A complete description of the sampling scheme and field measurement protocols is summarized by Campbell, Burrows, Gower, and Cohen (1999). Corn was the predominant cover type within this area, hence, a disproportionate
number (80%) of the plots were corn. On Day of Year (DOY) 253 in 1999, a destructive harvest was made to determine aboveground biomass. On this occasion, the density of plants was measured by direct counts of the number of stems per 5 m of row, and the number of rows per 25 m of field. One plant was then harvested from each quadrant of the $25 \times 25$-m plot. After drying, the average biomass per plant was scaled to the plot level by reference to the plant density. For the annual NPP estimates, biomass at DOY 253 was converted to NPP using an aboveground to total production ratio of 0.9 (Desjardins, 1985; Foth, 1962; Gower et al., 1999; Jones, Allen, Jones, Boote, & Campbell, 1984; Rochette, Desjardins, Pattey, & Lessard, 1995; Ruimy et al., 1994).

2.3. LUE by crop type

Determination of $e$, requires estimates of NPP and APAR. For the purposes of determining APAR, a continuous record of incident PAR and of $f_{\text{PAR}}$ is needed. Incident PAR data was available from a meteorological station less than 5 km from the study area maintained by the SURFRAD network (SURFRAD, 2000). Half-hourly average PAR values (W m$^{-2}$) were converted to joules and summed over the daylight periods to develop a daily time series for 1999 (Fig. 2). A corresponding $f_{\text{PAR}}$ time series for each crop type was derived from the LAI measurements.

The relationship of $f_{\text{PAR}}$ to LAI has previously been studied in corn and soybeans (Daughtry, Gallo, Gowar, Prince, & Kustas, 1992; Tollenaar & Bruulsema, 1988) and these relationships are of the form

$$f_{\text{PAR}} = \frac{a}{1 - e^{-k \times \text{LAI}}} \tag{1}$$

where $a$ = an empirical constant, $k$ = extinction coefficient, and LAI = projected leaf area per unit ground area.

For this study, we developed a representative LAI time series for each crop type using field measurements and estimated daily $f_{\text{PAR}}$ from these time series. Values for $a$ and $k$ (see Fig. 3) were taken from Daughtry et al. (1992).
The daily green LAI estimates were derived from LAI measurements in the early to mid-growing season at the same 80 plots used for NPP. At each of three sampling dates (DOY’s 147, 173, and 208), leaves were removed from one plant representing each quadrant of each plot and passed through a leaf area meter to determine projected green leaf area per plant. As with biomass, these values were scaled to the plot level by reference to the plant density. To create a continuous representative LAI time series for each crop type, mean LAI at each sampling date for each crop type was determined and an exponential function was fit to the means at the first two LAI measurement dates when leaf biomass was present (DOY’s 147 and 173 for corn and DOY’s 173 and 208 for soybean). The trajectories for the latter part of the growing season were based on the third sampling date and a green LAI time series reported at the eddy covariance flux tower (Ameriflux, 2001). There, continuous measurements of incoming and reflected PAR at the flux tower were used to develop daily LAI time series for corn in 1999 and soybean in 1998 (Tilden Meyers, National Oceanic and Atmospheric Administration, personal communication). Because the soybean LAI trajectory was from 1998, it may have differed somewhat from the actual 1999 trajectory. However, neither 1998 nor 1999 were extreme climate years (e.g., in both cases, annual precipitation was within 15% of the 30-year mean), so the trend at the end of the growing season was assumed to be similar. The final representative LAI trajectories (Fig. 3a) were consistent with earlier studies that have recorded the time course of LAI development in corn and soybean (Daughtry et al., 1992; Desjardins, 1985; Gallo, Daughtry, & Wiegand, 1993; Rochette et al., 1995). The LAI trajectories were converted to $f_{\text{APAR}}$ trajectories (Fig. 3b) by using Eq. (1). The product of daily incident PAR from the meteorological station (Fig. 2) and the daily $f_{\text{APAR}}$ yielded a daily estimate of APAR for each crop type. $e_a$ was estimated from the total NPP and the growing season APAR.

### 2.4. Area-wide NPP estimates

A 1-km resolution APAR time series was created to implement a LUE algorithm over the 25-km² study area. The 1999 AVHRR biweekly maximum value composite of the normalized difference vegetation index (NDVI) was used for the $f_{\text{APAR}}$ component (B. Reed, USGS EROS Data Center, personal communication). Theoretical studies have generally found a linear relationship between $f_{\text{APAR}}$ and satellite-based spectral vegetation indices such as NDVI, with minor variation across vegetation types (Sellers, 1987; Sellers et al., 1994). To create an NDVI/$f_{\text{APAR}}$ model specific to this study area, an estimate of $f_{\text{APAR}}$ at the 1-km resolution was made for each AVHRR grid cell in the 5 × 5-km study area for each 2-week AVHRR composite period. Direct spatial correspondence was achieved by overlaying the 25-m land cover grid with the 1-km AVHRR grid. The coarse resolution $f_{\text{APAR}}$ were derived by first assigning an $f_{\text{APAR}}$ to each ETM+ resolution cropland cell for each day based on the cover type, its associated $f_{\text{APAR}}$ time series (Fig. 3b), and the DOY. The cells in the Other class were assigned an $f_{\text{APAR}}$ of zero. Then the average of the 1600 $f_{\text{APAR}}$ within each 1-km cell was determined at the end of each 2-week compositing window. The end of the compositing window was used because it was most likely to match the maximum NDVI during the green-up phase of the growing season.

The time series of NDVI values for a particular cell in the AVHRR dataset is not a precise sample across time for a fixed 1-km grid cell on the ground because of the positional uncertainty of AVHRR imagery (±1 pixel) and the maximum value compositing procedure (Holben, 1986). The effect of this positional uncertainty is greatest when a 1-km area contains a significant proportion of the Other class with low NDVI. Because subgrid-scale areas of low NDVI tend not to be represented in the AVHRR time series, the three 1-km cells that had greater than 25% of their area in the Other class were not included in the development of the NDVI/$f_{\text{APAR}}$ regression.

The resulting relationship of NDVI to $f_{\text{APAR}}$ (Fig. 4) was used with the original NDVI time series over the complete growing season to generate the coarse resolution $f_{\text{APAR}}$ time series. The incident PAR time series was then applied to the biweekly $f_{\text{APAR}}$ to get the daily APAR time series at the 1-km resolution over the 25-km² study area.

To assess results from a coarse resolution LUE approach, NPPs were generated assuming the whole area had the $e_a$ of corn or the whole area had the $e_a$ of soybean. For a third
approach, the high spatial resolution land cover map was used to assign crop-specific $e_\alpha$. The Other class was assigned an $e_\alpha$ of zero for this case. A NPP reference data layer was created by summing the product of area and mean NPP, as measured by the plant samples, for the two crops. The NPP of the Other class was assumed to be zero.

3. Results

Within the 25-km$^2$ study area, 44.4% of the land was classified as corn, 44.4% as soybean, and 11.2% as Other. Classification accuracy of the 100 validation points was 94%. Small cropped areas that site visits had revealed as being neither corn nor soybean were misclassified as corn or soybean.

The LAI measurements indicated an earlier leaf-out and a higher maximum LAI for corn than soybean (Table 1). Differences between the crops in the $f_{\text{APAR}}$ time series were smaller than differences in the LAI time series (Fig. 3a,b). Corn reached an $f_{\text{APAR}}$ greater than 0.9 by DOY 177 and soybean followed 25 days later. Corn $f_{\text{APAR}}$ dropped below 0.9 on DOY 241 and soybean did so on DOY 252.

Total APAR on the biomass sampling date was 750 MJ m$^{-2}$ for corn and 601 MJ m$^{-2}$ for soybeans. Differences in APAR between the crop types were largest early in the growing season (Fig. 5) because of the earlier leaf-out of corn. The rate of APAR accumulation was similar between the two crop types once soybean was fully leafed out.

At the biomass sampling date, aboveground biomass for soybean was 55% of that for corn. Estimated NPP was 1180 (S.D. 218) g m$^{-2}$ for soybeans and 2159 (S.D. 558) g m$^{-2}$ for corn based on plant densities of 6.9 (S.D. 0.8) plants m$^{-2}$ for corn and 32.5 (S.D. 2.9) plants m$^{-2}$ for soybean. The $e_\alpha$ for the period up to DOY 253 was 2.88 g MJ$^{-1}$ for corn and 1.96 g MJ$^{-1}$ for soybean.

NDVI and $f_{\text{APAR}}$ increased in tandem over most of the growing season, then decreased late in the growing season (Fig. 4). The largest deviations between the observations and the least squares linear regression was for Periods 8 and 9 when there was a possible temporal mismatch between NDVI and $f_{\text{APAR}}$. The foliage was beginning to senesce during this period, with an associated increase in red reflectance but not near infrared reflectance—hence, a decrease in NDVI. The maximum value compositing procedure probably assigned a value from the beginning of the interval when the NDVI was relatively high. However, the green $f_{\text{APAR}}$ was from the end of the period when it had been reduced because of senescence.

For the reference NPP map based on the crop-specific areas and crop-specific NPPs, the average NPP for the 25-km$^2$ study area was 1482 g m$^{-2}$ and total biomass

---

**Table 1**

Green LAI measurements at three dates

<table>
<thead>
<tr>
<th>DOY</th>
<th>Corn (Mean)</th>
<th>S.D.</th>
<th>Soybeans (Mean)</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>147</td>
<td>0.07</td>
<td>0.03</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>173</td>
<td>3.42</td>
<td>1.14</td>
<td>0.38</td>
<td>0.08</td>
</tr>
<tr>
<td>208</td>
<td>7.68</td>
<td>1.47</td>
<td>4.08</td>
<td>0.96</td>
</tr>
</tbody>
</table>

---

**Fig. 5.** Cumulative APAR for corn and soybean.
production was \(3.70 \times 10^{10}\) g. With the LUE approach and the assumption of continuous corn, the estimated biomass production was 126% of the reference total (Table 2). For the set of cells that were actually corn, there was an average NPP difference of 286 g m\(^{-2}\) between the reference value and the LUE-based value. This effect was the result of the slightly lower \(f_{\text{APAR}}\) in the early growing season relative to what it would have been had the area actually been all corn. In contrast, the actual soybean cells were overestimated by an average of 693 g m\(^{-2}\) (a total of 0.77 \(\times 10^{10}\) g), predominantly because \(e_{n}\) was too high. In addition, the \(f_{\text{APAR}}\) early in the growing season was slightly higher than would be expected for an area that was uniformly soybean. The cells classified as Other were mapped as 1783 g m\(^{-2}\) (0.52 \(\times 10^{10}\) g). For the alternative assumption, in which the whole area was assigned a soybean \(e_{n}\), total biomass production was 86% of the NPP reference total. In this case, there were also compensating errors. The dominant factor was the assignment of a low \(e_{n}\) to corn grid cells causing an average underestimate of 878 g m\(^{-2}\) (0.97 \(\times 10^{10}\) g). The actual soybean cells were overestimated by an average of 101 g m\(^{-2}\) (0.11 \(\times 10^{10}\) g) because of the slightly higher \(f_{\text{APAR}}\) early in the growing season. The assumption of production by the Other cells resulted in an overestimate of 1281 g m\(^{-2}\) per cell (0.36 \(\times 10^{10}\) g). When the high-resolution land cover was used to prescribe \(e_{n}\), with \(f_{\text{APAR}}\) from the 1-km AVHRR, total biomass production was within 5% of the reference value.

Table 2
Average NPP and total biomass production for the study area using different algorithms

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>NPP (g m(^{-2}))</th>
<th>Total production (g (\times 10^{10}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crop-specific areas and NPPs</td>
<td>1482</td>
<td>3.70</td>
</tr>
<tr>
<td>LUE model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All corn</td>
<td>1873</td>
<td>4.68</td>
</tr>
<tr>
<td>All soybean</td>
<td>1281</td>
<td>3.20</td>
</tr>
<tr>
<td>High-resolution land cover</td>
<td>1402</td>
<td>3.51</td>
</tr>
</tbody>
</table>

and coarse resolution satellite imagery supports the general approach of monitoring green \(f_{\text{APAR}}\), and, hence, APAR, over large areas with coarse resolution remote sensing (Chihar, Chen, & Li, 1997).

Because of the positional uncertainty of the AVHRR sensor imagery, the biweekly compositing procedure used to generate large area NDVI surfaces is likely to produce an overestimate of average \(f_{\text{APAR}}\) when there is fine-scale heterogeneity in \(f_{\text{APAR}}\), as seen here. Pixel selection in the temporal compositing procedure is likely to avoid areas with a significant proportion of nonvegetated surface. Positional accuracy for the MODIS sensor products, which are being used in the MODIS Land Science Team LUE algorithm, is expected to be on the order of 0.1 pixel instead of one pixel as for AVHRR (Justice et al., 1998). MODIS-based estimates of \(f_{\text{APAR}}\) and APAR over heterogeneous areas are therefore likely to be an improvement over AVHRR-based estimates.

For a complete NPP monitoring algorithm, the other critical component besides APAR is the production efficiency factor, \(e_{n}\), which has been well studied in crop plants because of its potential as an indicator of growth potential (Kiniry et al., 1989; Sinclair & Horie, 1989) and because crops are a suitable model system for testing some of the assumptions of LUE-based NPP algorithms (Daugtry et al., 1992; Gallo et al., 1993). Results of these earlier studies support the result here that \(e_{n}\) can differ widely between crops, specifically corn and soybean (Daugtry et al., 1992; Gallo et al., 1985, 1993; Gower et al., 1999; Rochette et al., 1995). Thus, when corn and soybean are grown in close proximity, the assignment of an \(e_{n}\) to a coarse resolution pixel is problematic.

The mechanistic basis for the difference in \(e_{n}\) between corn and soybean relates in part to the differences in photosynthetic pathway: corn is a \(C_{4}\) species and soybean a \(C_{3}\) species. The leaf-level photosynthesis rate per unit APAR tends to be greater in \(C_{4}\) than \(C_{3}\) species under conditions of high light, drought, or high temperature (Ehleringer, 1978). The benefit of the \(C_{4}\) pathway becomes most significant at temperatures >25 °C. Since temperature often rises about 30 °C during the growing season in the Midwest, a differential sensitivity to temperature may in part account for the higher \(e_{n}\) for corn relative to soybean. Soybean may also have lower \(e_{n}\) because of energy expenditures to support nitrogen-fixing bacteria. In reviews of \(e_{n}\), Gower et al. (1999) and Prince (1991) found highest values in \(C_{4}\) species, intermediate values for non-N-fixing \(C_{3}\) species, and lowest values for N-fixing \(C_{3}\) species.

If corn and soybean were grown in fields that were large relative to the resolution of the satellite sensor used to monitor \(f_{\text{APAR}}\) in an operational LUE algorithm, and crop type was readily discernable by the same sensor, then an appropriate \(e_{n}\) could be assigned to each crop and production estimates could be made based on APAR and an associated land cover map. However, the principal implication of this study is that a LUE-based NPP algorithm run over the Midwest at coarse resolution (≥ 1 km) would be

4. Discussion

The linear relationship found in this study between average green \(f_{\text{APAR}}\) within 1-km cells and NDVI from the AVHRR sensor is consistent with similar relationships for several crop species found at the plot scale with boom-mounted sensors (Galoo, Daugtry, & Bauer, 1985; Gallo et al., 1993). The slope for the satellite-based relationship was 1.8 compared to 1.2 for a plot scale study with corn and soybean (Daugtry et al., 1992). Factors such as solar-surface-sensor geometry, atmospheric effects, background effects, and nongreen biomass (Huemmrich & Goward, 1997; Myneni, Asrar, Tanre, & Choudhury, 1992) contribute to the scatter in the satellite-based NDVI/\(f_{\text{APAR}}\) relationship, but a consistent signal across the growing season was apparent. This convergence of results from plot-level studies


MODIS Land Team (2001). Available at: http://mod08.gsfc.nasa.gov/MODIS/LAND/VAL/


Tol                                                 

